

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 014 282 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
28.06.2000 Bulletin 2000/26

(51) Int Cl. 7: G06F 17/30

(21) Application number: 99310164.1

(22) Date of filing: 17.12.1999

(84) Designated Contracting States:
 AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
 MC NL PT SE
 Designated Extension States:
 AL LT LV MK RO SI

(30) Priority: 22.12.1998 US 218434

(71) Applicant: Xerox Corporation
Rochester, New York 14644 (US)

(72) Inventors:
 • Williamowski, Jutta
 38100 Grenoble (FR)

- Pareschi, Remo
20122 Milano (IT)
- Borghoff, Uwe M.
38330 St. Ismier (FR)

(74) Representative: Skone James, Robert Edmund
 GILL JENNINGS & EVERY
 Broadgate House
 7 Eldon Street
 London EC2M 7LH (GB)

(54) Search channels between queries for use in an information retrieval system

(57) An information retrieval system (102-105) for retrieving information from multiple information sources. The information retrieval system provides for the building of dynamic queries through the use of query channels. A query channel permits the passing of attributes of the search results between different queries. The query channel can cause the automatic passing of the attributes, or it can be user controlled (breakpoints). Further, the query results may be transformed so that they are in the context of the target query (computational

transformations). For example, the results may be translated or reformatted into a form utilized by the target query. Query channels are created based on graphical representations of queries and their attributes. A simple drag and drop operation, wherein an attribute is selected and dragged to the target query, is used to create the channel. A channel may be created while a query is running. Specification of breakpoints or computational transformations are made using a pop-up window interface that may appear when the channel is selected.

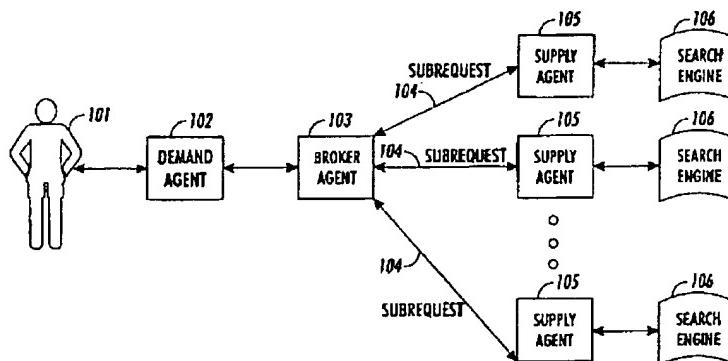


FIG. 1

EP 1 014 282 A1

1

EP 1 014 282 A1

2

Description

[0001] The present invention is related to the searching of information repositories such as databases, and in particular to a facility for generating and building dynamic queries.

[0002] When accessing heterogeneous legacy databases on Intranets or while querying multiple information sources on the Internet, the end-user only wants to have a simple and straightforward point of access.

[0003] An exemplary information retrieval architecture is described in the article entitled "System Components For Embedded Information Retrieval From Multiple Disparate Information Sources", Ramana B. Rao, Daniel M. Russell, and Jock D. Mackinlay, Proceedings of 1993 ACM Symposium on User Interface Software and Technology, Atlanta, GA, November 1993 ACM SIGGRAPH and SIGCHI. The architecture incorporates an intermediary server which mediates access requests between an information access client (i.e. the user) and various information sources. Thus, the user only needs to interface with the information access client in order to retrieve the information from multiple information sources.

[0004] Another difficulty in information retrieval is query generation. The process of generating an effective query is often an iterative one. In such a process, the results of one query are then used to formulate a second, often more narrow query. This occurs because the first query may have been too broad in scope and had returned large amounts of unusable or redundant data. Moreover, known query systems are "static" in nature. In other words, in order for query to begin, all of the query terms must be fixed. However, because of the vast amount of information constantly becoming available on the World Wide Web, it would be desirable to make queries more dynamic in nature. Further, since the information repositories are disparate, portions of the query may be running in parallel. Thus, it would be desirable to be able to specify and execute a query that is dynamic in nature so that search terms are not static.

[0005] The present invention is related to the searching of network accessible distributed databases, such as those found on the Internet. Drag and drop search channels are described which enable the combining of queries. A query is visually displayed on the display of a computer system as a query window. The query window will provide for the input of attributes that define the nature and scope of intended query. The attributes may be logically combined to create the intended query. Each attribute specification is represented as an object. In order to include an attribute of a first query as an attribute of a second query, a user selects an attribute, wherein a graphical representation of the attribute is created. The user then "drags and drops" the graphical representation of the attribute to an attribute specification area of the second query. Thus, a channel of communication between two queries is established which can

have the properties described below.

[0006] The Drag and Drop Search channels provide Computational Transformations. Computational Transformations allows for the automatic adaptation of the results associated with a channel to the query context.

[0007] The Drag and Drop Search channels further provide for the ability to create compound queries dynamically and incrementally and further provide breakpoints so that search results may be operated on and manually filtered by the user.

5 have the properties described below.
[0006] The Drag and Drop Search channels provide Computational Transformations. Computational Transformations allows for the automatic adaptation of the results associated with a channel to the query context.

[0007] The Drag and Drop Search channels further provide for the ability to create compound queries dynamically and incrementally and further provide breakpoints so that search results may be operated on and manually filtered by the user.

10 Figure 1 is a block diagram illustrating the components of and process flow for an information retrieval system of the currently preferred embodiment of the present invention;

15 Figures 2 and 3 illustrate simple and complex query windows as may be utilized in the currently preferred embodiment of the present invention;

20 Figures 4 and 5 illustrate simple and complex result windows as may be utilized in the currently preferred embodiment of the present invention;

25 Figures 6 and 7 illustrate by way of sample screen displays, an example of the creation of channels between queries, as may be performed in the currently preferred embodiment of the present invention;

30 Figures 8 and 9 illustrate by way of sample screen displays, a user interface for specifying computational transformations or breakpoints in a channel between queries, as may be performed in the currently preferred embodiment of the present invention;

35 Figures 10 and 11 illustrate the data structure for a complex query which is used to show the channel relationship between queries, as may be used in the currently preferred embodiment of the present invention; and,

40 Figure 10 is an illustration of a computer-based system as may be used to implement the currently preferred embodiment of the present invention.

45 [0008] The present invention is directed to an interface for query generation to and information retrieval from one or more information sources. The currently preferred embodiment is implemented in an information retrieval system which provides software agents that can query multiple heterogeneous databases, reconstruct information and generate customized reports. A description of the system is provided below, while a more comprehensive description as of the time of filing this application can be found on the World Wide Web at URL <http://www.xrce.xerox.com/research/ct/research/cbkb.html>.

50 [0009] The present invention builds on the system by providing a facility to build dynamic complex queries. A dynamic complex query is one where query attribute specification is dynamic, rather than static, so that the results (or partial results) of one query can be used as an attribute specification by another query. Furthermore, the facility is made available to a user in a very

intuitive fashion using well-known drag and drop graphical user interface methods.

[0010] The currently preferred embodiment of the present invention can be implemented for use on computer based systems capable of accessing the World Web (the Web) using a suitable Java™ enabled Web Browser. The various software programs utilized in the implementation are written in the well-known Java programming language.

[0011] The amount of information available from electronic sources on the World Wide Web or other on-line information repositories is increasing exponentially. Tools are needed to extract relevant information from these repositories. The information retrieval system sophisticated facilities for efficient information retrieval and combination. Figure 1 illustrates how information requests are processed in the information retrieval system. Referring to figure 1, a user 101 requests information through a Demand Agent, 102. The Demand Agent 102 typically embodies a User Interface which is presented to a User. In the currently preferred embodiment, the Demand Agent 102 is a Java™ Applet which is downloaded to User's personal computer system and is accessible using a suitably configured Web Browser (such as Netscape Navigator® having the appropriate Java™ software capabilities)

[0012] Demand Agents 102 will then present the information request to a Broker Agent 103. The Broker Agents 103 decompose information requests into inter-dependent subrequests 104. The requests or subrequests that cannot be further decomposed are fed into Supply Agents 105 that access electronic information repositories through search engines 106. The Supply Agents 105 will typically be configured to interface using the protocols of a particular search engine 106. For example, a Supply Agent may be provided to interface and make requests to the Alta Vista™ search engine. In the currently preferred embodiment, the Broker Agents and Supply Agents will operate on a Server system which provides information retrieval.

[0013] The results of the information request are received and then recomposed by the Broker Agent 103. It should be noted that the information retrieval system will provide for advanced querying facilities for the user not provided by some information sources, so the re-composition includes further filtering of the results to enable the advanced querying facilities. The recomposition may further include such tasks as formatting the received information and/or ranking the results with respect to the query request. Once the recomposition is completed, the results are returned to the Demand Agents 102 for presentation to the user 101.

[0014] It should be noted that preferably, requests, information and results are represented via feature constraints. Requests do not need to be fully defined; they may correspond to partial specifications of the requested information. Furthermore, requests that cannot be fully satisfied may still obtain results in the form of partial

objects that refine the initial requests by instantiating some of its attributes or by adding new attributes, thus providing user feedback.

[0015] Various aspects of the information retrieval system can be found in the following PCT Publications WO 98/48359 (29 October 1998) and WO 98/48361 (29 October 1998).

[0016] Each of the various agents described above are preferably embodied as software programs, written in the Java programming language for execution on suitably configured computer based systems.

[0017] Figure 2 is an illustration of a query window for specifying a query. Referring to Figure 2, a domain area 201 of query window 200 provides a means for specifying one of a plurality of predetermined domains for the query. The query domain relates to the type of information being sought. For example, a query domain may be a type of documents (e.g. books or articles) or a subject area (e.g. physics or operas). In the currently preferred embodiment, these domains are predefined and correspond to one or more information sources to which the query will be directed. The areas 202-204 provide for specifying a search pattern. The area 202 is for specifying a predetermined attribute (e.g. title or author). The area 203 is for specifying a constraint (e.g. contains or matches exactly) which is to be placed on the attribute and a specified keyword. The area 204 is for specifying a desired keyword(s), which may define an expression formed by keywords and Boolean operators. If the search pattern is correct, pressing the button 205 labeled "Add to the Specification" causes the pattern to be added to the current specification of the query, as indicated in specification area 206. If the search pattern is not correct, it can be reset using a reset button 207.

[0018] The specification area 206 includes all the desired search patterns specified by the user.

[0019] Buttons 208-211 provide for editing the specification. Button 208 provides for editing a selected search pattern. Button 209 provides means for removing a search pattern. Button 210 provides means for clearing the entire specification. Button 211 provides for cancellation of a sub-query.

[0020] Finally, buttons 212-213, respectively, provide for the control of submission and cancellation of a query. Figure 3 illustrates a window for performing a sub-query within a query, i.e. a complex query. In a complex query, a subquery generates a set of subresults. The results generated are subresults in the sense that they deliver values that will be used in the query pattern of the main query (or in general a higher level query). A sub-query is provided for by selecting a predetermined option in the area 203, namely the "answer of" option. This causes the window of Figure 3 to be displayed. As can be seen in figure 3, a new Tab 301 has been included to indicate that it is a sub-query. This also enables the user to traverse between the main query and the sub-query. Note that the rest of the sub-query window is similar to the main query.

[0020] Figures 4 and 5 illustrate summary and detail views of the result window from a search. Referring to Figure 4, a summary view shows a list of all the hits and some overall summary information such as the number of hits. In Figure 5, all information fields relating to a particular "hit" are displayed. A body area 501 shows the textual content of the document. Information area 502 provides various information such as date, URL location, information source, size and title. In this view, traversal through the list of hits is provided previous 503 and next 504 buttons provided within the window.

[0021] Establishing a Drag-and-Drop Channel between two queries amounts to an instruction to pass answers, returned for an attribute field of the source query, as input values for an attribute field in the target query. In one respect, this is conceptually analogous to the relationship between a query and a sub-query. However, the Drag and Drop Channel of the present invention, provides further functionality. The Drag-and-Drop Channels provide :

- living links between queries which allow the user to construct compound queries dynamically and incrementally,
- computational transformations which allow the user to automatically adapt the dragged-and-dropped attribute values to the query context, and
- breakpoints for user intervention into the search process which allow the user to filter the values passed along the Drag-and-Drop Channels.

[0022] Living links between queries allow the user to construct compound queries dynamically and incrementally. The user can introduce Drag-and-Drop Channels between queries not only *a priori*, but also *a posteriori* with respect to launching the involved queries.

[0023] Usually, e.g. with Structured Query Language (SQL), the definition of compound queries must be done *a priori*. The user must know in advance how they will link several queries to obtain what they are searching for so the user defines all involved queries and links them before launching the whole compound query.

[0024] But this is not always the natural way to create compound queries. Often the inspection of query results leads to new ideas on how to use them. But, using already obtained query results as input for another query is only possible, if queries can be combined dynamically. This is possible through Drag-and-Drop Channels. They can be established *a posteriori*, when the source query is already (and continuously) running. This *a posteriori* operation is enabled as a result of a feature provided in the information retrieval system wherein query results may be continually accumulated (i.e. query results may be continuously returned as a result of new documents being placed within the query domain).

[0025] If the Drag-and-Drop Channel is introduced *a posteriori*, the already present results are immediately forwarded to the target query. In any case, all the results

arriving after introduction of the Drag-and-Drop Channel are forwarded as they arrive.

[0026] Computational Transformations provide for the automatic adaptation of the dragged-and-dropped attribute values to the query context. Passing the values from the source attribute of the Drag-and-Drop Channel to its target attribute is not always immediately possible. Problems arise, e.g. if the source and target attributes do not have the same type, arity or unit, or if the general context of the linked queries is different. Then computational transformations are necessary to enable the value passing between the queries. Below are examples of computational transformations:

15 Example (attribute type integration)

[0027] If you drop an integer attribute on to a string attribute, you have to transform the integer values of the output attribute field to a string value for the input attribute field.

Example (attribute arity integration)

[0028] If the arity of the linked attributes is not the same, especially if the source attribute of the Drag-and-Drop Channel contains a *list of values* and the target attribute expects *single values*, then these single values must be extracted from the list before they can be passed to the target query. E.g. if you drop an attribute whose value contains name *lists* (composed for example of all co-authors of found articles) on an attribute which expects *single names* (e.g. to search for the person's homepage), then you first have to extract the elementary (sur-)names of the authors, and pass them one-by-one to the input attribute field.

Example (attribute unit integration)

[0029] If the units of the source and the target attribute are different, the values to drop can be automatically calculated, e.g. multiply by 1000 to transform 1 km into 1 000 m, or divide by 60 to transform 1 minute into 1/60 h.

Example (linguistic integration)

[0030] The transformations can also be determined by other characteristics of the linked queries, such as their language. If for each query the language is defined (or can be inferred by language guessing for example), and two queries in different languages are linked, then the transformation consists for example in translating the values from the language of the source query to the language of the target query. A manner of performing such cross-lingual information retrieval is described in a copending U.S. Patent Application entitled: "System For Providing Cross-Lingual Information Retrieval", serial no. Not Yet Assigned.

[0031] If the characteristics of the source and target

attributes and queries is known by the system, then these computational transformations can be automatically introduced. Otherwise they can be user defined.

[0032] Drag-and-Drop Channels constitute possible breakpoints for user intervention into the search process. In fact, often the user does not want to automatically pass all the values found for the source attribute of the Drag-and-Drop Channel to its target attribute. Instead, the user wants to control this process and to filter the values manually. Therefore a Drag-and-Drop Channel can be automatic or user-controlled. The user can specify for each Drag-and-Drop Channel if the values shall pass automatically or if the user wants to filter them.

[0033] This is especially useful if the query results of the source query have low precision, i.e. if many of the documents in the query result are not adequate for use as input into a subsequent query. If all inadequate values are automatically passed to the subsequent searches, the final results of the compound query would get a very low precision. Thus, making a Drag-and-Drop Channel user-controlled, allows the user to improve the precision of the final results. Furthermore it improves the efficiency of the search process as it also reduces the number of queries generated by the passed values.

[0034] The usage of Drag-and-Drop Channels as user filtering points extends also the functionality of a Drag-and-Drop Channel introduced *a posteriori*: before establishing the Drag-and-Drop Channel by dragging the source attribute onto the target attribute, the user might want to suppress some inadequate results or attribute values. This corresponds to a *posteriori* filtering. In consequence the afterwards established Drag-and-Drop Channel is a user-controlled one.

[0035] Figures 6-7 are illustrative screen displays indicating how channels between queries are created. Referring to Figure 6, the "author" attribute 611 of query 1 601 has been linked to query 4 604. Here, the author attribute 612 of query 2 602 is also to be linked as an attribute of query 4. As indicated by the highlighting, the author attribute has been selected and is being moved via a drag and drop user interface operation. Figure 7 illustrates the results wherein query 4 604 now also contains the results of the author attribute 612 of query 2 602 are included as an attribute of query 4 604.

[0036] Figure 8 is an illustrative screen display showing an interface for specifying operational characteristics of a channel between queries. Referring to Figure 8, assuming that the channel with query 2 has been selected and a user interface action, such as a double mouse click, has been performed to cause a pop-up control window 800 be displayed. This pop-up control window is used to select the various control features for a channel. So for example, a first selection area 801 has check boxes 801 and 802 which permits one to select between automatic or user controlled transfer of attribute query results. This is the manner in which breakpoints can be introduced in the target query (in this case query 4).

[0037] A second selection area 806 has a check box 804 that is used to specify if it is desirable to automatically transform the attribute results to be in the target query context. Here, the transformation is automatically identified 805 based on the source and target query context (e.g. query 2 presents retrieves Author information in List format, whereas query 4 retrieves Author information in single name format). The above is but one way of how operational characteristics of a link can be specified and is not meant to be limit the scope of the present invention. Other ways could be utilized, such as through a control checkbox associated with a channel, and would not cause departure from the spirit and scope of the present invention.

[0038] It should be noted that the user interfaces illustrated in Figures 6-8 would typically be presented to a user through a Demand Agent, as illustrated in Figure 1.

[0039] Figure 9 is an example of a data structure for a complex query showing the relationship of channels between queries. It should be noted that a simple query would typically have a structure found in one of the sub-queries described in Figure 9. In any event, referring to Figure 9, lines 901-905 define the parts of the query. In this case, three (3) sub-queries are defined along with two (2) computational transformations. The definition of query 1 begins at line 906 and ends at line 916. From these lines it is shown that Query 1 has a domain of "computer science", a query attribute wherein the title contains the string "database systems" and a query attribute of author which is filled in by the query. A computational transformation of type "nameExtraction", is performed on the content of the author attribute and stored in a location indicated as "A11" (lines 911-916).

[0040] The definition of query 2 begins at line 917. The domain of query 2 is "computer science" and the title should have the word "constraints" in it. In query 2, the author attribute is subject to the same computational transformation as found in query 1 Lines 922-926).

[0041] Finally, the query 3 begins on line 927. In query 3, the domain is "computer science" and the author attribute contains the results of the author attributes of queries 1 and 2. This is the channel established between query 3 and each of queries 1 and 2.

[0042] Typically, the data structure of Figure 9 would be interpreted by a broker agent as illustrated in Figure 1.

[0043] The computer-based system on which the currently preferred embodiment of the present invention may be implemented is described with reference to Figure 10. For example, the computer based system 10 may embody a user system upon which a demand agent executes and the user interface presented to a user, or it may be server system wherein the various broker and supply agents would execute. In any event, referring to Figure 10, the computer-based system is comprised of a plurality of components coupled via a bus 1001. The bus 1001 illustrated here is simplified in order not to obscure the present invention. The bus 1001 may consist

of a plurality of parallel buses (e.g. address, data and status buses) as well as a hierarchy of buses (e.g. a processor bus, a local bus and an I/O bus). In any event, the computer system is further comprised of a processor 1002 for executing instructions provided via bus 1001 from Internal memory 1003 (note that the Internal memory 1003 is typically a combination of Random Access or Read Only Memories). The processor 1002 and Internal memory ROM 1003 may be discrete components or a single integrated device such as an Application Specification Integrated Circuit (ASIC) chip which has been designed to embody the functional features of the present invention.

[0044] Also coupled to the bus 1001 are a keyboard 1004 for entering alphanumeric input, external storage 1005 for storing data, a cursor control device 1006 for manipulating a cursor, and a display 1007 for displaying visual output. The keyboard 1004 would typically be a standard QWERTY keyboard but may also be telephone like keypad. The external storage 1005 may be fixed or removable magnetic or optical disk drive. The cursor control device 1006 will typically have a button or switch associated with it to which the performance of certain functions can be programmed. The network connection 1008 would be used to receive and/or transmit data containing a bit-mapped representation of a medium.

Claims

1. An information retrieval system for generating queries to one or more information sources, said information retrieval system comprising:

a demand agent (102) providing:

a user interface for allowing a user to specify a search expression;
a query specification interface for allowing a user to specify a channel between a first query and a second query wherein the search results for an attribute of said first query are provided as an attribute of said second query;

a broker agent (103) providing

a search and retrieval element for processing a search expression and creating specific information source search requests; a channel element for associating the results of an attribute of said first query with said second query; and
a supply agent (105) for receiving said search requests and requesting and obtaining documents from information sources responsive to said search requests.

2. The system as recited in claim 1, wherein said demand agent (102) is further comprised of a channel control interface for causing user controls on said channel, wherein a first user control is a computational transformation for transforming the results of the search for said first attribute into the query context of said second query; and wherein said broker agent (103) is further comprised of an element for causing an associated computational transformation to occur.

3. The system as recited in claim 2, wherein said computational transformation is one of a translation from a source language to a target query language, an attribute type integration, an attribute arity integration, and an attribute unit integration.

4. A method for obtaining information from one or more information sources from an information retrieval system, said method comprising the steps of:

- a) a user specifying one or more query attributes for a first query;
- b) said user specifying a second query, wherein in at least one of said attributes for said second query is the results of a query attribute from said first query;
- c) said information retrieval system providing the results of said first query;
- d) said information retrieval system extracting from the results of said first query, the results for said specified query attribute;
- e) said information retrieval system providing said results for said specified query attribute as a result for said second query; and
- f) said information retrieval system providing said user with the results of said second query.

5. A program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform method steps for obtaining information from one or more information sources from an information retrieval system, said method comprising the steps of:

- a) a user specifying one or more query attributes for a first query;
- b) said user specifying a second query, wherein in at least one of said attributes for said second query is the results of a query attribute from said first query;
- c) said information retrieval system providing the results of said first query;
- d) said information retrieval system extracting from the results of said first query, the results for said specified query attribute;
- e) said information retrieval system providing said results for said specified query attribute as

11

EP 1 014 282 A1

12

a result for said second query; and
f) said information retrieval system providing
said user with the results of said second query.

5**10****15****20****25****30****35****40****45****50****55****7**

EP 1 014 282 A1

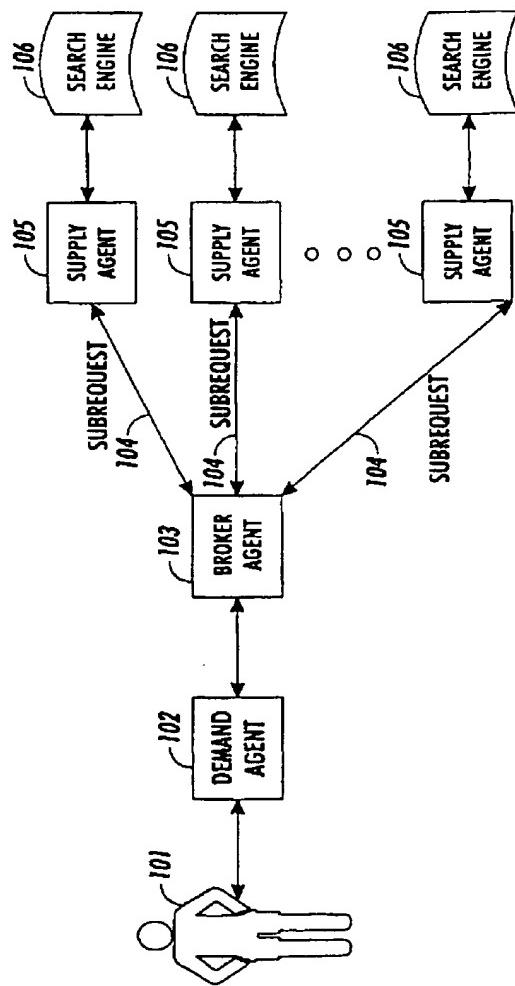


FIG. 1

EP 1 014 282 A1

MAINQUERY

203

200

201

202

204

207

205

206

208

209

210

211

212

213

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

EP 1 014 282 A1

301

MAINQUERY	SUBQUERY 1	
QUERY CONCERNING OPERAS		
COMPOSER	CONTAINS	WAGNER
RESET ADD TO SPECIFICATION		
EDIT REMOVE CLEAR CANCEL SUBQUERY		
GLOBAL OPTIONS:		<input type="checkbox"/> CASESENSITIVE SEARCH <input type="checkbox"/> ALSO SEARCH WEBCRAWLERS
		SUBMIT CANCEL

© 1998 XEROX CORPORATION" (17 U.S.C. 401)

FIG. 3

EP 1 014 282 A1

RESULTS OF QUERY 1 (BOOKS/ARTICLES) : DATE AFTER 9506 ...

SHOW	SUMMARY	USING	TITLE	NB OF HITS: 24
<p>HIT 1: GENERAL AGENTS WEB SERVICE MAINPAGE HIT 2: IWAH - INTELLIGENT WEB AGENTS / HOUSTON HIT 3: PERSISTENT JAVA AND MOBILE WEB AGENTS HIT 4: REAL ESTATE AGENTS WEB PAGES CREATED BY INNOVATIVE... HIT 5: INTERNET AGENTS: SPIDERS, WANDERERS, BROKERS, AND 'BOTS/FAH-CHENG HIT 6: PROGRAMMING INTELLIGENT AGENTS FOR THE INTERNET/MARK WATSON HIT 7: GENERAL AGENTS WEB SERVICES MAINPAGE HIT 8: INTERACTIVE HOME BUYING ON WEB - REAL ESTATE AGENTS & COMPANIES HIT 9: HTML WEB PAGE DESIGN AND MARKETING AGENTS HIT 10: ETHICAL WEB AGENTS HIT 11: TAHITIWEB - TAHITI AND ITS ISLANDS WEB PAGES - TRAVEL AGENTS HIT 12: WWW-VRML OCT-DEC 1995: INTERACTIVE AGENTS AND THE WEB HIT 13: WWW-VRML OCT-DEC 1995: RE: INTERACTIVE AGENTS AND THE WEB HIT 14: DARTFLOW: A WORKFLOW MANAGEMENT SYSTEM ON THE WEB USING TRANSP HIT 15: TRAVEL_AGENTS INTERNET MAILING LIST HOT 16: INTERNET AGENTS (BOOK OUTLINE)</p>				
<p>ALPHABETIC SORT QUALITY SORT</p> <p>ALL TO HTML SELECTED TO HTML DELETE SELECTED HITS UNSELECT ALL</p>				

"© 1998 XEROX CORPORATION" (17 U.S.C. 401)

FIG. 4

EP 1 014 282 A1

RESULTS OF QUERY 1 (BOOKS/ARTICLES): TITLE CONTAINS WEB OR INTERNET ...

SHOW ALL FIELDS OF MAINQUERY NB OF HITS: 25

HIT 8

BODY

501

WELCOME TO THE GENERAL AGENTS WEB SERVICES
MAINPAGE. WE ARE A FULL-SERVICE WEB DESIGN AND
HOSTING SERVICE. PLEASE CHOOSE ONE OF THE
FOLLOWING LINKS FOR ...

DATE 09.05.96

HTTP_URL [HTTP://WWW.GENERALAGENTS.COM/](http://WWW.GENERALAGENTS.COM/)

INFORMATION_SOURCE ALTA VISTA (DIGITAL)

SIZE 2K

TITLE GENERAL AGENTS WEB SERVICES MAINPAGE

503 504

PREVIOUS ... NEXT ... GO TO: QUALITY SORT

ALL TO HTML SELECTED TO HTML DELETE SELECTED HITS UNSELECT ALL

"© 1998 XEROX CORPORATION" (17 U.S.C. 401)

FIG. 5

EP 1 014 282 A1

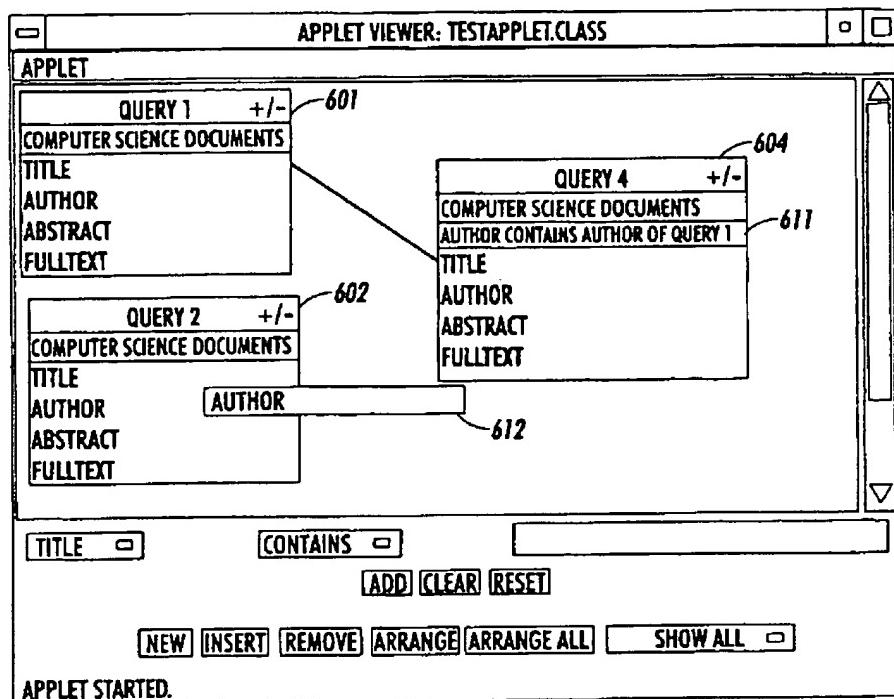
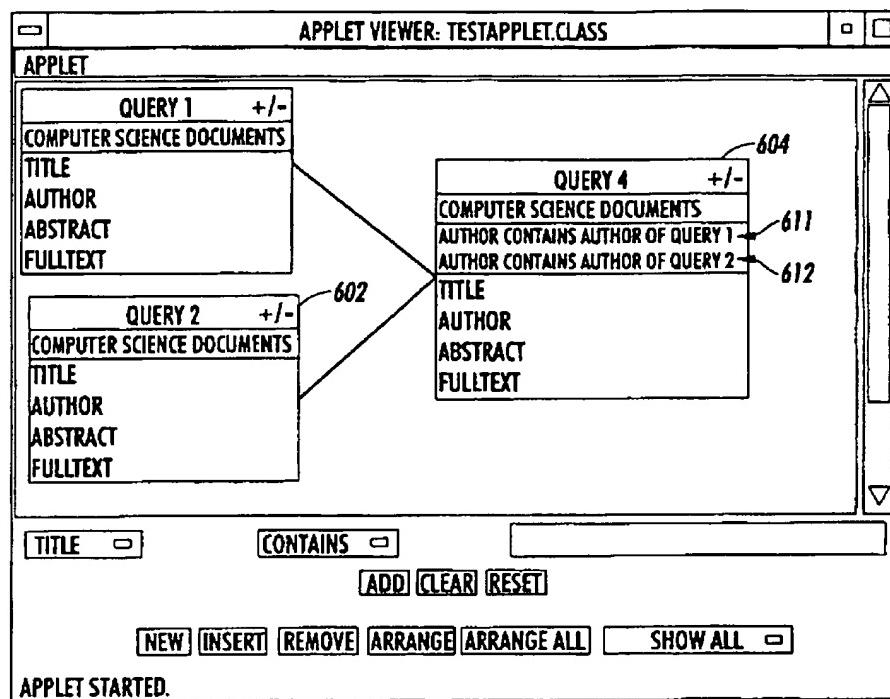


FIG. 6

EP 1 014 282 A1



"© 1998 XEROX CORPORATION" (17 U.S.C. 401)

FIG. 7

EP 1 014 282 A1

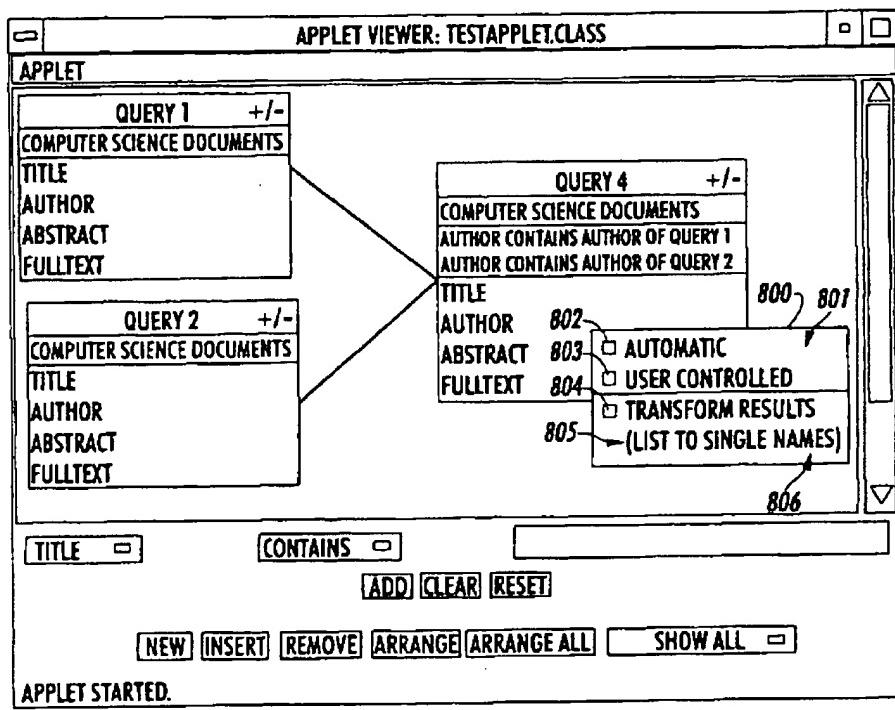


FIG. 8

EP 1 014 282 A1

```

901 CompQuery: part1 -> QUERY1           //CompoundQuery contains a query QUERY1
902 CompQuery: part2 -> QUERY2           //CompoundQuery contains a query QUERY2
903 CompQuery: part3 -> QUERY3           //CompoundQuery contains a query QUERY3
904 CompQuery: part4 -> CT1              //CompoundQuery contains a computational transformation CT1
905 CompQuery: part5 -> CT2              //CompoundQuery contains a computational transformation CT2
906 QUERY1: domain -> D1                //QUERY1 has a domain D1
907 D1 == "computer science"            //which is "computer science"
908 QUERY1: title -> T1                //QUERY1 has a title T1
909 T1 contains TC1                  //whose value contains [should contain]
910 TC1 == "database systems"          //the string "database systems"
911 QUERY1: author -> A1              //QUERY1 has also an author A1
912                                         // which will be filled by the query results
913 CT1: domain -> CTD1              // the computational transformation CT1
914 CTD1 == "nameExtraction"          //is of type nameExtraction
915 CT1: nameList -> A1              //it will extract all names from A1
916 CT1: name -> A11                // and provide the values as result in A11
917 QUERY2: domain -> D2              //A11 is used as input in QUERY3
918 D2 == "computer science"          //A11 is used as input in QUERY3
919 QUERY2: title -> T2              //A11 is used as input in QUERY3
920 T2 contains TC2                //A11 is used as input in QUERY3
921 TC2 == "constraints"             //A11 is used as input in QUERY3
922 QUERY2: author -> A2              //A11 is used as input in QUERY3
923 CT2: domain -> CTD2              //A11 is used as input in QUERY3
924 CTD2 == "nameExtraction"          //A11 is used as input in QUERY3
925 CT2: nameList -> A2              //A11 is used as input in QUERY3
926 CT2: name -> A21                //A11 is used as input in QUERY3
927 QUERY3: domain -> D3              //A11 is used as input in QUERY3
928 D3 == "computer science"          //A11 is used as input in QUERY3
929 QUERY3: author -> A3              //A11 is used as input in QUERY3
930 A3 contains A11                //A11 is used as input in QUERY3
931 A3 contains A21

```

FIG. 9

EP 1 014 282 A1

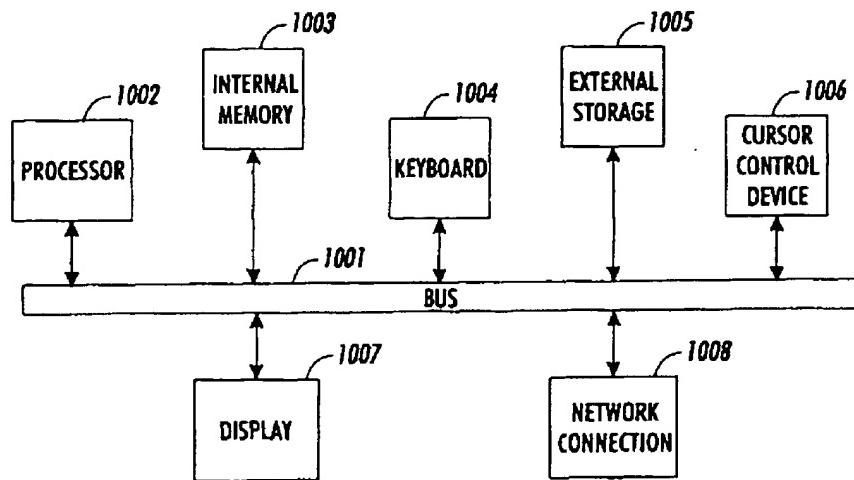


FIG. 10

EP 1 014 282 A1



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 31 0164

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int.Cl.)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
A	EP 8 747 845 A (IBM) 11 December 1996 (1996-12-11) * column 21, line 14 - column 22, line 9 * * column 15, line 25 - column 16, line 43 * * column 12, line 25 - line 48 * * column 6, line 3 - column 7, line 2; figures 1,7,9,11 *	1,4,5	G06F17/30
A	WO 98 29792 A (INTEL CORP) 9 July 1998 (1998-07-09) * page 15, line 10 - page 17, line 6; figure 5 * * page 11, line 13 - page 13, line 13; figure 3 *	1,4,5	
A	CHEN J R ET AL: "COLLABORATIVE INFORMATION AGENTS ON THE WORLD WIDE WEB" ACM CONFERENCE ON DIGITAL LIBRARIES, US, NEW YORK, NY: ACM, 1998, pages 279-280. XP000853973 ISBN: 0-89791-965-3 * the whole document *	1,4,5	TECHNICAL FIELDS SEARCHED (Int.Cl.)
A	CHEN L. ET AL.: "Multi-Agents Based Dynamic Request Placement Strategies in Distributed Information Systems" PROC. 9TH. INT. CONF. ON PARALLEL AND DISTRIBUTED COMPUTING SYSTEMS, PDCS '96, vol. 2, 25 - 27 September 1996, pages 701-708, XP0008892307 Dijon, France * page 706, right-hand column, line 1 - line 24; figure 8 *	1,4,5	G06F
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
BERLIN	17 March 2000	Deane, E	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons A : technological background O : non-written disclosure P : intermediate document R : member of the same patent family, corresponding document	
EPO FORM 1501 (02-92) (PCT/01)			

EP 1 014 282 A1**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 99 31 0164

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
 The members are as contained in the European Patent Office EDP file on
 The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

17-03-2000

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0747845	A	11-12-1996	US 5721908 A BR 9602597 A CA 2177917 A CN 1138177 A JP 9026972 A	24-02-1998 06-10-1998 08-12-1996 18-12-1996 28-01-1997
WO 9829792	A	09-07-1998	US 5963944 A AU 5594998 A DE 19782227 T GB 2335770 A	05-10-1999 31-07-1998 18-11-1999 29-09-1999

EPO FORM P01B
For more details about this annex : see Official Journal of the European Patent Office, No. 12/82